

RECEIVED

NOV 26 2004

Technology Center 2600

Attorney Docket No. JP919990203



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

**Patent Application**

Applicant(s): Y. Tahara et al.  
Docket No.: JP919990203  
Serial No.: 09/656,963  
Filing Date: September 7, 2000  
Group: 2655  
Examiner: Michael N. Opsasnick

I hereby certify that this paper is being deposited on this date with the U.S. Postal Service as first class mail addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450

Signature:

*Lisa L. Tulpis*

Date: November 19, 2004

Title: Methods and Apparatus for Voice Information Registration  
and Recognized Sentence Specification in Accordance  
With Speech Recognition

---

**SECOND SUPPLEMENTAL APPEAL BRIEF**

Mail Stop Appeal Brief - Patents  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

This Second Supplemental Appeal Brief is submitted in response to the Office Action dated August 19, 2004 in the above-referenced application, in which the Examiner reopened prosecution in response to the Supplemental Appeal Brief filed May 25, 2004.

Appellants have submitted concurrently herewith a response to the Office Action, requesting reinstatement of the appeal pursuant to 37 C.F.R. §1.193(b)(2).

**REAL PARTY IN INTEREST**

The present application is assigned to International Business Machines Corp., as evidenced by an assignment recorded December 5, 2000 in the U.S. Patent and Trademark Office at Reel 11315, Frame 0734. The assignee, International Business Machines Corp., is the real party in interest.

## RELATED APPEALS AND INTERFERENCES

There are no known related appeals and interferences.

## STATUS OF CLAIMS

Claims 1-15 are pending in the present application. Claims 5, 10 and 15 are allowable, and claims 1-4, 6-9 and 11-14 stand rejected under 35 U.S.C. §103(a). Claims 1-4, 6-9 and 11-14 are appealed.

## STATUS OF AMENDMENTS

There have been no amendments filed subsequent to the rejection.

## SUMMARY OF INVENTION

The present invention relates to a word registration method for a speech recognition system and, more particularly, to a method whereby voice is used to specify information displayed on a screen (Specification, page 1, lines 5-8).

By way of example, as recited in claim 1, a voice information registration method, employed by a speech recognition apparatus, comprises the following steps. A sentence group is obtained, which includes a first to an  $N$ -th sentence, wherein  $N$  is a number equal to or greater than two. A sounds-like spelling is obtained for a word that is included in an  $i$ -th sentence, but is not entered in a speech recognition dictionary, wherein  $i$  is a number equal to or less than  $N$ . A base form is obtained based on the sounds-like spelling of the word. Finally, the base form is registered in a speech recognition dictionary in correlation with the word.

In an illustrative embodiment, a group of sentences to be recognized is obtained from an application. Using parsing logic, each target sentence to be recognized is divided into words, or speech recognition units. Thereafter, the words in each target sentence are examined to determine whether there are unknown words among them that are not registered in the speech recognition dictionary, but for which the sounds-like spelling is available. If an unknown word is found, a base form, from which the pronunciation is inferred from the sounds-like spelling, is prepared and

registered in the speech recognition dictionary. This base form is employed when the voice of a user is recognized who has orally designated one of the sentences (Specification, page 2, lines 4-12)

According to one aspect of the present invention, a voice information registration method is provided, which is employed by a speech recognition apparatus, and with which a voice input device is used (Specification, page 2, lines 13-15). According to another aspect of the present invention, a sentence specification method is provided that is employed by a speech recognition apparatus, and with which a voice input device is used. This sentence specification method has both a registration step and a recognition step (Specification, page 2, lines 24-26).

An unknown word, detected in an extracted sentence, is recognized as a word but is not registered in the speech recognition dictionary. Thus the base form of the unknown word is unknown to the system (Specification, page 17, lines 5-7). A flow diagram showing the unknown word detection processing performed according to an embodiment of the present invention is shown in FIG. 5. A flow diagram showing the processing performed according to an embodiment of the present invention to obtain a base form corresponding to an unknown word and to register the base form in a speech recognition dictionary is shown in FIG. 6. As described in the figures, the present invention allows for the recognition of a sentence that includes words that are not registered in a speech recognition dictionary through the registration process of the present invention (Specification, page 26, lines 10-12).

#### ISSUES PRESENTED FOR REVIEW

1. Whether claims 1-4, 6-9 and 11-14 are properly rejected under 35 U.S.C. §103(a) as being unpatentable over U.S. Patent No. 5,208,897 to Hutchins (hereinafter "Hutchins"), in view of U.S. Patent No. 6,233,553 to Contolini et al., and IBM Technical Disclosure Bulletin, Vol. 35, Issue 1a, p. 59 (hereinafter "IBM TDB").

#### GROUPING OF CLAIMS

Claims 1-4, 6-9 and 11-14 do not stand or fall together. More particularly, claims 1, 6 and 11 stand or fall together and claims 2-4, 7-9 and 12-14 stand or fall together.

## ARGUMENT

Appellants incorporate by reference herein the disclosure of all previous responses filed in the present application, namely, Amendment and Response to Office Action dated February 19, 2003, Response to Final Office Action dated July 31, 2003, Appeal Brief dated October 6, 2003 and Supplemental Appeal Brief dated May 25, 2004.

With regard to the issue of whether claims 1-4, 6-9 and 11-14 are properly rejected under 35 U.S.C. §103(a), Appellants respectfully reassert that the cited combination fails to establish a prima facie case of obviousness under 35 U.S.C. §103(a), as specified in M.P.E.P. §2143.

As set forth therein, M.P.E.P. §2143 states that three requirements must be met to establish a prima facie case of obviousness. First, the cited combination must teach or suggest all the claim limitations. Second, there must be some suggestion or motivation to combine reference teachings. Third, there must be a reasonable expectation of success. While it is sufficient to show that a prima facie case of obviousness has not been established by showing that one of the requirements has not been met, Appellants respectfully believe that none of the requirements have been met.

First, there is a clear lack of motivation to combine the references. For at least this reason, a prima facie case of obviousness has not been established. Hutchins is directed to the performance of speech recognition, while Contolini is directed to the generation of a phonetic transcription associated with a spelled word and IBM TDB is directed to building word models using sounds-like spellings. That is, the teachings in each reference are directed to completely different processes in speech recognition technology; one (Hutchins) toward actual real-time recognition of a spoken utterance, the other two (Contolini and IBM TDB) toward building models that may eventually be used in actual real-time recognition of a spoken utterance. However, other than a very general and conclusory statement in the Office Action, there is nothing in the three references that reasonably suggests why one would actually combine the teachings of these three references.

The Federal Circuit has stated that when patentability turns on the question of obviousness, the obviousness determination “must be based on objective evidence of record” and that “this precedent has been reinforced in myriad decisions, and cannot be dispensed with.” In re Lee, 277 F.3d 1338, 1343 (Fed. Cir. 2002). Moreover, the Federal Circuit has stated that “conclusory statements” by an examiner fail to adequately address the factual question of motivation, which is

material to patentability and cannot be resolved “on subjective belief and unknown authority.” *Id.* at 1343-1344.

In the Office Action at page 3, paragraph 3, the Examiner provides the following statement to prove motivation to combine Hutchins and Contolini, with emphasis supplied: “[t]herefore, it would have been obvious . . . to modify the teachings of Hutchins (5208897) with a trained sound dictionary of words as specified by the user because it would advantageously allow for a higher recognition rate tailored for individual users and more efficient dictionaries.”

Additionally, in the Office Action at page 4, paragraph 1, the Examiner provides the following statement to prove motivation to combine Hutchins, Contolini and IBM TDB, with emphasis supplied: “[t]herefore, it would have been obvious . . . to modify the teachings of Hutchins in view of Contolini (6233553) with a ‘sounds like spelling’ technique because it would advantageously allow user to enter the information more accurately than the phonetic pronunciations.”

Since Hutchins has no disclosure relating to word registration, it is unclear why there would be any motivation to combine Hutchins with Contolini or IBM TDB. Although the Examiner cites “IBM TDB, disclosure text, near the end,” after the above statement, Appellants submit that this statement is based on the type of “subjective belief and unknown authority” that the Federal Circuit has indicated provides insufficient support for an obviousness rejection. More specifically, the Examiner fails to identify any objective evidence of record which supports the proposed combination.

Second, Appellants assert that there is no reasonable expectation of success in achieving the present invention through a combination of Hutchins, Contolini and IBM TDB. For at least this reason, a prima facie case of obviousness has not been established. Despite the assertion in the Office Action, Appellants do not believe that Hutchins, Contolini and IBM TDB are combinable since it is not clear how one would combine them. There is no guidance provided in the present Office Action. However, even if combined, for the sake of argument, they would not achieve the automated registration techniques of the claimed invention.

Third, Appellants assert that even if combined, the combination fails to teach or suggest all of the limitations of the claims. For at least this reason, a prima facie case of obviousness has not been established.

By way of further explanation, the present specification, at page 1, lines 8-19, describes one problem that the claimed invention addresses:

As is described in Japanese Unexamined Patent Publication No. Hei 10-320168, the disclosure of which is incorporated by reference herein, a conventional method is available whereby voice is used to specify information displayed on a screen. However, to use this method, a menu or a button in an application, and a sentence in which a link to a web is included must be registered using words that can be recognized by a speech recognition system.

All of the character strings for a menu, in this case, can be statically added to a speech recognition dictionary, but since the web link would tend to be changed daily, coping with such a change would exceed the capabilities of a method for which static registration is employed. In addition, if too many words, more than are necessary, are added to the dictionary, other problems, such as a reduction in the recognition accuracy or an extended processing time, may be encountered.

The Office Action, at page 3, paragraph 2, contends that Hutchins discloses the steps/operations of independent claims 1, 2, 6, 7, 11 and 12 including: “obtaining a sentence group . . . sentence;” “obtaining a spelling . . . speech recognition dictionary;” “obtaining a base form . . . word;” and “registering said base form . . . said word.” The Examiner acknowledges that Hutchins fails to disclose the concept of adding to a dictionary and the use of sounds-like spellings, however, points to Contolini and IBM TDB in providing these aspects. The Office Action then summarily concludes that it would have been obvious to combine the three references to achieve the claimed invention. Appellants strongly disagree. All of Hutchins’ operations are associated with the actual recognition of speech uttered by a user. The techniques of Hutchins do not relate to word registration in a speech recognition dictionary, as recited in the elements of claims 1, 2, 6, 7, 11 and 12.

The first step of independent claim 1 of the present invention recites “obtaining a sentence group. . .” The Office Action contends that a section of Hutchins describing word grammar (words to phrases) discloses this step. However, this section of Hutchins discloses the process of combining

individual words from digitized speech to form phrases in the word recognition process. This combination of individual words does not disclose the obtaining of groups of sentences in a word registration process.

The second step of independent claim 1 of the present invention recites “obtaining a sounds-like spelling for a word that is included in an *i*-th sentence, but is not entered in a speech recognition dictionary . . .” The Examiner admits that Hutchins does not disclose the use of a sounds-like spelling. However, the Examiner contends that a section of Hutchins describing the ASCII spelling of words as output resulting from incoming speech, discloses the remainder of this step. However, ASCII output of Hutchins occurs once the word recognition process is complete and the spoken words were found in a recognition dictionary. Therefore, Hutchins does not disclose the providing a spelling of a word that is not entered in a speech recognition dictionary as part of a word registration process.

The third step of independent claim 1 of the present invention recites “obtaining a base form based on said sounds-like spelling of said word . . .” The Office Action alleges that a section of Hutchins that relates to processing subsyllables and syllables discloses this step. However, in the present invention the base form of a word is based on a sounds-like spelling in a word registration process. Further, the base form is from a word not registered in the speech recognition dictionary. This differs significantly from syllable processing in a speech recognition process, as described in Hutchins.

The fourth step of independent claim 1 of the present invention recites “registering said base form in a speech recognition dictionary in correlation with said word.” The Office Action again alleges that a section of Hutchins relating to a word recognition process combining subsyllables to form syllables, and combining syllables to form words, discloses this step. These syllables and subsyllables are not being registered as a part of the speech registration process. Thus, Hutchins does not disclose the speech registration elements recited in independent claim 1 or the remaining independent claims of the present invention.

The Examiner acknowledges that Hutchins teaches away from adding to a dictionary since Hutchins does not relate to a word registration process. However, the Examiner cites Contolini in an attempt to incorporate continued adding to a dictionary. Contolini discloses an automatic

determination of phonetic transcriptions associated with spelled words. A tool allows a user to specify a confidence level, and a system stores all generated pronunciations that fulfill the confidence level, updating a pronunciation dictionary or lexicon.

The combination of Hutchins and Contolini fails to remedy all of the deficiencies of Hutchins presented above. For example, the combination fails to disclose the obtaining of a sentence group. As admitted by the Examiner, Contolini contains no disclosure of sounds-like spellings and only discloses the conversion of a spelled word into a phonetic transcription. Finally, the combination fails to disclose the registering of the base form of a sounds-like spelling in a speech recognition dictionary. Contolini contains no disclosure of a base form of a word or the registration of a base form from a pronunciation dictionary in a speech recognition dictionary.

Furthermore, with respect to IBM TDB, while discussing the use of “sounds-like spellings,” IBM TDB explains that a user may enter a “sounds-like spelling” to assist in building word models. This significantly differs from the steps/operations of the claimed invention since IBM TDB, other than stating that a user may enter a “sounds-like spelling,” is silent as how a word model is formed given the “sounds-like spelling.” Thus, IBM TDB suggests nothing about obtaining a sentence group, obtaining a sounds-like spelling for a word that is included in an *i*-th sentence, but is not entered in a speech recognition dictionary, obtaining a base form based on said sounds-like spelling of said word, and registering said base form in a speech recognition dictionary in correlation with said word, as recited in the claimed invention.

Appellants do not assert that they have developed the concept of “sounds-like spellings.” Such spellings are known, for example, as evidenced by IBM TDB. However, the automated use of “sounds-like spellings” as recited in the registration techniques of the claimed invention was not known prior to the invention and is clearly not taught or suggested by the combination of Hutchins, Contolini and IBM TDB.

Also, with specific regard to claims 2, 7 and 12, despite a contention in the Office Action to the contrary, the combination of Hutchins, Contolini and IBM TDB fails to disclose obtaining voice information that is input as a user reads and vocally reproduces a display corresponding to the *i*-th sentence, as in the claimed invention.



Appellants point out that while their arguments may, at times, discuss Hutchins, Contolini and IBM TDB one at a time, it is in an effort to clearly illustrate that the references do not teach or suggest one or more of the elements of the claimed invention. Thus, in general, by pointing out that reference 1 fails to disclose an element, say element A, and then pointing out that references 2 and 3 also fails to disclose element A, the discussion effectively points out that the combination of the three references, even if proper, would fail to disclose element A.

Therefore, for at least the reasons given above, Appellants again respectfully request that the §103(a) rejections of independent claims 1, 2, 6, 7, 11 and 12 be withdrawn.

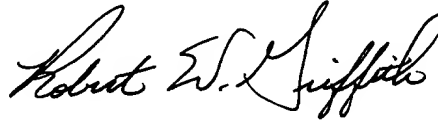
Furthermore, it is respectfully reasserted that the claims which respectively depend from independent claims 2, 7 and 12, i.e., claims 3, 4, 8, 9, 13 and 14, are patentable over the cited combinations. Thus, Appellants again request withdrawal of the §103(a) rejections of said claims.

In addition, it is asserted that dependent claims 3, 4, 8, 9, 13 and 14 recite patentable subject matter in their own rights. Claims 3, 8 and 13 recite that the group of sentences is obtained from an application and that a control message corresponding to the *i*-th sentence is generated and transmitted to the application. There is no application or control message disclosed in Hutchins, Contolini or IBM TDB. Further, claims 4, 9 and 14 recite that a sounds-like spelling score is stored in correlation with the sounds-like spelling of the word, that a pronunciation score is stored in correlation with the base form, and that the base form is registered in a speech recognition dictionary when a function value that is obtained by using the sounds-like spelling score and the pronunciation score exceeds a threshold value. Again, Hutchins, Contolini and IBM TDB are completely silent as to the use of any scores stored in correlation with sounds-like spellings and base forms. In response to the arguments previously presented by the Appellants, the Examiner stated that the response fails to comply with 37 C.F.R. §1.111(b) because they amount to a general allegation that the claims define a patentable invention without specifically pointing out how the language of the claims patentably distinguishes them from the references. However, as presented above, the patentable language of the claims is fully presented and not disclosed in the references cited by the Examiner.

For at least the reasons given above, Appellants respectfully request withdrawal of the §103(a) rejections of claims 1-4, 6-9 and 11-14. Appellants believe that claims 1-4, 6-9 and 11-14

are not obvious in view of Hutchins, Contolini and IBM TDB. As such, the application is asserted to be in condition for allowance, and favorable action is respectfully solicited.

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Robert W. Griffith". The signature is fluid and cursive, with the first name "Robert" and last name "Griffith" being clearly legible.

Date: November 19, 2004

Robert W. Griffith  
Attorney for Applicant(s)  
Reg. No. 48,956  
Ryan, Mason & Lewis, LLP  
90 Forest Avenue  
Locust Valley, NY 11560  
(516) 759-4547